

Application No. 10/044,711

Atty Docket: INXT 1018-1

**REMARKS**

Claims 1-34 are pending in this application. Claims 12, 14, 24 and 26 have been indicated by the Examiner as having allowable subject matter and if rewritten to overcome certain rejections and contain certain limitations, would be allowed. The remaining claims have various rejections as itemized further below. Many of the claims have been amended.

Applicants greatly appreciate the Examiner's sincere and detailed efforts to present the positions asserted. Welcome to the Examining Corps!

**The Drawings**

FIGs. 5A-5B drawings are objected to because the Examiner does not understand them and, particularly, because there are not reference numbers or a legend. Applicants believe that paragraph [0047] on pages 13-14 already provides sufficient information for one of skill in the art to understand the figures and, therefore, traverses the objection. The relevant paragraph describes the relationship of a test or probe document to its nearest neighbors and the importance of normalizing metrics. Clearly, the squares in the middle of FIGs. 5A-5B are the test documents. As the application explains:

These formulas may be modified to reflect the number of documents among the nearest neighbors that share the topic assignment  $T_m$ . Figures 5A-B illustrate the significance of normalizing scores. An ambiguous situation, as in Figure 5A, may appear to strongly support a topic assignment, because the neighborhood of the test document is dense. An unambiguous situation, as in Figure 5B, may seem to weakly support a topic assignment, unless a normalized metric of confidence is used, because the neighborhood to the test document is less dense.

One concept referenced in paragraph [0047] and illustrated by FIGs. 5A-5B is density. Density is the number of things in an area or within a metric radius. In this case, density is the number of circles or triangles within a radius of the test document. FIG. 5A is the "denser", because the circles and triangles are closer, within a shorter radius of the test document than in FIG. 5B. In the words of the application, FIG. 5A is more ambiguous than FIG. 5B, because the nearest neighbors of the test document belong evenly to two sets, the circles set and the triangles set. FIG. 5B is unambiguous, because the nearest

**Application No. 10/044,711****Atty Docket: INXT 1018-1**

neighbors predominantly belong to the triangles set.

Applicants respectfully submit that one of ordinary skill in the art would readily understand FIGs. 5A-5B without any changes. The Examiner has not cited any evidence to the contrary. On the other hand, with above explanation in mind, if the Examiner would suggest how the figures should be annotated without introducing new matter, Applicants almost certainly would make the suggested annotations or changes. Of course, the description above is now part of this prosecution history.

### **The Specification**

The first paragraph of the application, showing related application information, has been amended.

The remaining objections to the specification, as specifically delineated in Page 4 of the Office Action, have been dealt with and the corrections have been made.

### **Claim Objections**

Claim 1 has been objected to because of the redundancy in the wording. The claim has been amended to overcome this objection.

### **Claim Rejections under 35 USC 112**

Claims 1-21 and 23-24 are rejected under 35 USC 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which Applicants regard as the invention. Particular rejections are addressed in the order presented by the Examiner.

The Examiner rejected claim 1 because it was unclear to the Examiner how to read the phrase "reporting the test documents and category assignments". Applicants note that "category assignments" is used in claim 1 on line 10 exclusively in the context of test documents. Applicants do not understand the source of the Examiner's confusion. Nonetheless and without narrowing the claim, we have amended claim 1, line 11 to parrot the language in line 10.

Application No. 10/044,711

Atty Docket: INXT 1018-1

The Examiner rejected claims 5-9 and 27 on the basis that the word "substantially" is a relative term and asserted that one of ordinary skill in the art would not be reasonably apprised of the scope of the invention. As the preface to section 2173.05 of the MPEP says, "The fact that claim language, including terms of degree, may not be precise, does not automatically render the claim indefinite under 35 U.S.C. 112, second paragraph. *Seattle Box Co., v. Industrial Crating & Packing, Inc.*, 731 F.2d 818, 221 USPQ 568 (Fed. Cir. 1984). Acceptability of the claim language depends on whether one of ordinary skill in the art would understand what is claimed, in light of the specification." When the Examiner quoted MPEP 2173.05(b), one of the balancing statements of the MPEP was omitted. "The court held that the limitation 'which produces substantially equal E and H plane illumination patterns' was definite because one of ordinary skill in the art would know what was meant by 'substantially equal.' *Andrew Corp. v. Gabriel Electronics*, 847 F.2d 819, 6 USPQ2d 2010 (Fed. Cir. 1988)." With the balance taught in the MPEP more clearly in mind, we turn to the particular terms "substantially all" and "substantially without".

Unlike the term "substantially increase" in *In re Nehrenberg*, the term in claim 5 is "substantially all." The specification contrasts this term in paragraph [0040] against less than exhaustive testing at 10-50 percent levels. Accordingly, one of ordinary skill in the art would understand what is claimed and the section 112 rejection should be withdrawn.

The term "substantially without user intervention" is used in claims 6-9 and 27. Applicants submit that this phrase provides the greatest clarity and precision that can be provided to capture the sense that the user sets the process in motion and does not have to do much, if anything, for the partitioning, categorizing and calculating steps to proceed. This satisfies the standard of MPEP 2173.05(a). In addition, one of ordinary skill could consult the manual for the CATEGORIZER™ for an illustration of "substantially without user intervention." That manual is part of the record in this application, because it is part of the provisional application that is incorporated by reference. "If the proposed language is not considered as precise as the subject matter permits, the examiner should provide reasons to support the conclusion of indefiniteness and is encouraged to suggest alternatives that are free from objection." *Id.* In the absence of reasons that "substantially without user intervention" would be

**Application No. 10/044,711****Atty Docket: INXT 1018-1**

misunderstood and lacking suggested alternatives, Applicants respectfully submit that these rejections should be withdrawn.

The Examiner rejected claims 11 and 13, 23 and 25 and the claims that depend from them on the basis that the antecedent basis for "the test document" is the plural form, "the test documents." Applicants doubt that this wording would cause any confusion. Nonetheless, without narrowing the claims, the terms "a particular test document" and "the particular test document" have been substituted for the singular "test document". This amendment should overcome the rejections.

The Examiner correctly rejected claims 15 and 28 and 29, 30 and 34 and the claims that depend from them on the basis of lack of antecedent basis for "the identifying step" or "the identified documents." The step identified should be the "reporting step", so the claims have been amended to overcome the rejections.

The Examiner correctly rejected claims 18-20 and 31-33 on the basis of lack of antecedent basis for "the user interface." The claims have been amended to overcome the rejections.

The Examiner rejected claims 29 and 30 on the basis of lack of antecedent basis for "the filtering step." The antecedent basis appears in claim 28, line 2. In addition, the reference has been shortened to "filtering".

Applicants note that the Examiner has disavowed having made a complete list of bases for rejection under section 112. Applicants trust that the Examiner will not make final any new rejection that could have been responded to in this paper if it had been raised.

#### **Claim Rejections under 35 USC 102**

Claims 22, 23 and 27 are rejected under 35 USC 102(b) as being anticipated by USP 5,251,131 to Masand et al. Applicants have amended claim 22, for which claims 23 and 27, to emphasize auditing a training set, which is not discussed in Masand et al. The amended claim provides:

Application No. 10/044,711

Atty Docket: INXT 1018-1

*A computer assisted method of auditing a superset of training data, the superset comprising examples of documents having one or more preexisting category assignments, the method including:*

*determining k nearest neighbors of the documents in a test subset automatically partitioned from the superset;*

*automatically categorizing the documents based on the k nearest neighbors into a plurality of categories;*

*calculating a metric of confidence based on results of the categorizing step and comparing the automatic category assignments for the documents to the preexisting category assignments; and*

*reporting the documents and preexisting category assignments that are suspicious and the automatic category assignments that appear to be missing from the documents, based on the metric of confidence.*

We point out that the adjectives "preexisting" and "automatic" applied to category assignments are added for ease of reference and not to narrow the claim. Applying these adjectives to the reporting step clarifies and somewhat broadens that step.

The cited reference does not address auditing training data. The passage cited discusses traditional application of classification methods to a "new sample". See, e.g., Col. 23, lines 21, 26; col. 24, line 11. The invention, as a whole, for auditing training data and identifying mistakes in the training data is not taught or suggested by Masand et al.

Applicants respectfully submit that claim 22 and claims 23 and 27 that depend from it are allowable over Masand et al. for at least the reasons stated.

### Claim Rejections under 35 USC 103

Claims 1, 5-11, 15-18, 20 and 21 are rejected under 35 USC 103(a) as being unpatentable over USP 6,301,579 to Becker, and further in view of USP 5,251,131 to Masand et al. As neither Becker nor Masand et al. teach or suggest auditing training data and identifying mistakes in the training data, the combination cannot solve a problem that neither of them individually mentions, addresses, teaches or suggests.

Applicants have amended claim 1, from which claims 5-11, 15-18, 20 and 21 depend, to emphasize auditing a training set, which is not discussed in Becker or Masand et al.

The amended claim provides:

Application No. 10/044,711

Atty Docket: INXT 1018-1

*A computer assisted method of auditing a superset of training data, the superset comprising examples of documents having one or more preexisting category assignments, the method including:*

*partitioning the superset into at least two disjoint sets, including a test set and a training set, wherein the test set includes one or more test documents and the training set includes examples of documents belonging to at least two categories;*

*automatically categorizing the test documents using the training set;*

*calculating a metric of confidence based on results of the categorizing step and comparing the automatic category assignments for the test documents to the preexisting category assignments; and*

*reporting the test documents and preexisting category assignments that are suspicious and the automatic category assignments that appear to be missing from the test documents, based on the metric of confidence.*

As with claim 22, the adjectives "preexisting" and "automatic" applied to category assignments are added for ease of reference and not to narrow the claim. Applying these adjectives to the reporting step clarifies and somewhat broadens that step.

The Becker reference does not describe auditing a training set to identify mistakes, it illustrates classifier rules derived from the training set. It calculates probability estimates that a new sample will be given a particular classification under particular conditions (col. 29, lines 50-51), rather than calculating a metric of confidence related to whether preexisting category assignments in a test set are wrong. That is part of what compels the Examiner to concede that "Becker fails to teach reporting documents and category assignments that are *suspicious* and that *appear to be missing* based on the *metric of confidence*." Becker does not address the same problem and does not calculate the same metrics so there is no chance that Becker would report mistakes in the training data.

In a footnote, the Examiner "asserts that visualizing a decision tree classifier (col. 29, lines 37039) is performed by reporting test set data." Applicants do not see support for this assertion in Becker's figures. Figures 10B and 22 are interesting visualizations of the decision tree classifier. Neither of them report mistakes in training data. Overall, Becker has nothing to do reporting mistakes in the training data.

**Application No. 10/044,711****Atty Docket: INXT 1018-1**

The Examiner proposes to combine Becker with Masand et al. to add a new reporting feature that is described by Masand et al. at col. 22, line 33 – col. 24, line 44. Reading Masand et al. makes it clear that the feature being added is reporting potential mistakes in assignment of *new samples*. See, e.g., Col. 23, lines 21, 26; col. 24, line 11. Nothing in Masand et al. teaches or suggests detecting mistakes in a training set. Combining Becker with Masand et al. cannot supply the elements that are missing from each of them individually.

Applicants respectfully submit that claims 1, 5-11, 15-18, 20 and 21 are allowable over Masand et al. for at least the reasons stated.

**Claims 2 and 4 are rejected under 35 USC 103(a) as being unpatentable over USP 6,301,579 to Becker in view of USP 5,251,131 to Masand and further in view of USP 6,324,531 to Anderson et al.**

Claim 3 is rejected under 35 USC 13(a) as being unpatentable over USP 6,301,579 to Becker in view of USP 5,251,131 TO Masand in view of USP 6,324,531 to Anderson and further in view of USP 5,537,488 issued to Menon et al.

Claim 13 is rejected under 35 USC 103(a) as being unpatentable over USP 6,301,579 to Becker in view of USP 5,251,131 to Masand and further in view of 6,643,629 to Ramaswamy et al.

Claim 19 is rejected under 35 USC 103(a) as being unpatentable over USP 6,301,579 to Becker in view of USP 5,251,131 to Masand and further in view of USP 6,405,195 to Ahlberg et al.

Claim 25 is rejected under 35 USC 103(a) as being unpatentable over USP 5,251,131 to Masand in view of USP 6,643,629 TO Ramaswamy et al.

Claims 28-31 and 33-34 are rejected under 35 USC 103(a) as being unpatentable over USP 5,251,131 to Masand in view of USP 6,301,579 to Becker.

Claim 32 is rejected under 35 USC 103(a) as being unpatentable over USP 5,251,131 to Masand in view of US 6,405,195 to Ahlberg.

Applicants respectfully submit that claims 2 and 4, 3, 13, and 19 are allowable for at least the same reasons as claim 1. Claims 25, 28-31 and 33-34 and 32 are allowable for at least the same reasons as claim 22.

**Application No. 10/044,711**

**Atty Docket: INXT 1018-1**

**CONCLUSION**

Applicants respectfully submit that the pending claims are now in condition for allowance and thereby solicit acceptance of the claims, in light of these amendments.

The undersigned can ordinarily be reached at his office at (650) 712-0340 from 8:30 to 5:30 PST, M-F and can be reached at his cell phone (415) 902-6112 most other times.

Respectfully submitted,

  
\_\_\_\_\_  
Ernest J. Beffel, Jr., Reg. No. 43,489

Dated: 17 December 2004

Haynes Beffel & Wolfeld LLP  
P.O. Box 366  
Half Moon Bay, CA 94019  
(650) 712-0340 (telephone)  
(650) 712-0263 (facsimile)